

RAMSEY GROWTH IN SOME NIP STRUCTURES

ARTEM CHERNIKOV, SERGEI STARCHENKO, AND MARGARET E. M. THOMAS

ABSTRACT. We investigate bounds in Ramsey theorem for relations definable in NIP structures. We generalize a theorem of Bukh and Matousek [6] from the semialgebraic case to arbitrary polynomially bounded o-minimal expansions of \mathbb{R} , and show that it doesn't hold in \mathbb{R}_{exp} . We also prove an analog for relations definable in the field of p -adics. Generalizing [11], we show that in distal structures the upper bound for k -ary definable relations is given by the exponential tower of height $k - 1$.

1. INTRODUCTION

We recall a fundamental theorem of Ramsey.

Let X be a set and let $E \subseteq X^k$ be a k -ary relation on X . We say that a sequence $(a_i : 1 \leq i \leq m)$ of elements in X is *E -indiscernible* (also called “ E -homogeneous” in the literature) if either E holds on all k -tuples $(a_{i_1}, \dots, a_{i_k})$ with $1 \leq i_1 < \dots < i_k \leq m$, or E doesn't hold on any $(a_{i_1}, \dots, a_{i_k})$ with $1 \leq i_1 < \dots < i_k \leq m$.

Fact 1.1. [32] *For every $k, n \in \mathbb{N}$ there is some number $N \in \mathbb{N}$ such that if X is a set and $E \subseteq X^k$ is a k -ary relation on X , then every sequence of elements of X of length N contains an E -indiscernible subsequence of length n .*

We denote the smallest such N by $R_k(n)$.

Establishing exact bounds for the asymptotics of $R_k(n)$ is one of the central open problems in combinatorics, even in the case $k = 2$. We summarize briefly some of the known results.

- Fact 1.2.** (1) [15, 18] $2^{\frac{n}{2}} < R_2(n) < 2^{2^n}$ for all $n > 2$.
 (2) [16, 17] There are positive constants c and c' such that $2^{cn^2} < R_3(n) < 2^{2^{c'n}}$ for all sufficiently large n .
 (3) [12, 19] For each $k \geq 3$ there are positive constants c, c' such that $\text{twr}_{k-1}(cn^2) \leq R_k(n) \leq \text{twr}_k(c'n)$ for all sufficiently large n , where the tower function $\text{twr}_k(n)$ is defined recursively by $\text{twr}_1(n) = n$ and $\text{twr}_{i+1}(n) = 2^{\text{twr}_i(n)}$.

Recently, this question was investigated in the context of semialgebraic relations, where stronger bounds were obtained. Recall that a set $A \subseteq \mathbb{R}^d$ is *semialgebraic* if it is given by a finite Boolean combination of sets of the form $\{x \in \mathbb{R}^d : f(x) \geq 0\}$, where $f(x)$ is a polynomial in d variables with coefficients in \mathbb{R} . We say that a semialgebraic set A has *description complexity* at most t if $d \leq t$ and A can be

A.C. was supported by the NSF Research Grant DMS-1600796 and by an Alfred P. Sloan Fellowship.

S.S. was supported by the NSF Research Grant DMS-1500671.

M.T. was supported by the DFG Research Grant TH 1781/2-1 and by the Zukunftskolleg, University of Konstanz.

written as a Boolean combination involving at most t different polynomials, each of degree at most t .

Definition 1.3. Let $E \subseteq (\mathbb{R}^d)^k$ be a k -ary semialgebraic relation on \mathbb{R}^d . For $n \in \mathbb{N}$, we let $R_E(n)$ be the smallest natural number N such that if $(a_i : 1 \leq i \leq m)$, $a_i \in \mathbb{R}^d$ is a sequence of length $m \geq N$, then it contains an E -indiscernible subsequence of length n .

Let $R_k^{d,t}(n)$ be the maximum of $R_E(n)$, where E varies over all k -ary semialgebraic relations on \mathbb{R}^d of description complexity at most t .

The case of binary relations ($k = 2$) is addressed in the following theorem, which shows that $R_2^{d,t}(n)$ can be bounded by a polynomial in n — as opposed to the necessarily exponential bound in the general case (Fact 1.2(1)). This essentially corresponds to the Erdős-Hajnal conjecture for semialgebraic graphs.

Fact 1.4. [2] *For any d, t there is some $c = c(d, t)$ such that $R_2^{d,t}(n) \leq n^c$ for all sufficiently large n .*

Based on it, [11] addresses the case of general k , establishing that $R_k^{d,t}(n)$ can be bounded from above by an exponential tower of height $k - 1$ (as opposed to k for general relations, Fact 1.2(3)).

Fact 1.5. [11] *For any $k \geq 2$ and $d, t \geq 1$ there is some $c = c(k, d, t)$ such that $R_k^{d,t}(n) \leq \text{twr}_{k-1}(n^c)$ for all sufficiently large n .*

Besides, matching lower bounds were obtained, demonstrating that the stepping-up lemma can be carried out “semialgebraically”.

Fact 1.6. (1) [11] *For every $k \geq 4$, there exists $d = d(k)$, $t = t(k)$, $c' = c'(k)$ and a k -ary semialgebraic relation E on \mathbb{R}^d of description complexity $\leq t$ such that $R_E(n) \geq \text{twr}_{k-1}(c'n)$ for all sufficiently large n .*
 (2) [13] *In (1), one can take $d = k - 3$.*

The dependence of the dimension d on the arity k of the relation E in Fact 1.6 is unavoidable, due to the following theorem of Bukh and Matousek.

Fact 1.7. [6] *For every $k \in \mathbb{N}$ and every k -ary semialgebraic relation E on \mathbb{R} there is some $c = c(E)$ such that $R_E(n) \leq 2^{2^{c^n}}$ for all sufficiently large n .*

That is, if we restrict to arbitrary k -ary semialgebraic relations on \mathbb{R} (as opposed to \mathbb{R}^d for some $d > 1$), then $R_E(n)$ is at most double exponential (rather than a tower of height $k - 1$ as in Fact 1.5). The constant c here given by their proof actually depends on the parameters of E (and not just on its description complexity, as in Fact 1.5).

In this paper we investigate a generalization from semialgebraic relations to relations definable in arbitrary first-order structures, and the connection between Ramsey growth for relations definable in a structure and the model-theoretic tameness conditions that this structure satisfies.

Definition 1.8. Let \mathcal{M} be a first-order structure in a language \mathcal{L} (we denote by M its underlying set). Let $\varphi(x_1, \dots, x_k)$ be an $\mathcal{L}(M)$ -formula (i.e. a formula with parameters from \mathcal{M}) with its free variables partitioned into k groups of equal size: $|x_1| = \dots = |x_k| = d$. Then φ defines a k -ary relation $\varphi(M)$ on M^d , namely $\varphi(M) = \{(a_1, \dots, a_k) \in (M^d)^k : \mathcal{M} \models \varphi(a_1, \dots, a_k)\}$.

We let $R_\varphi(n)$ be the smallest natural number N such that any sequence $(a_i : 1 \leq i \leq N), a_i \in M^d$ of length N contains a $\varphi(M)$ -indiscernible subsequence of length n .

Also, given an \mathcal{L} -formula $\varphi(x_1, \dots, x_k; z)$, where $|x_1| = \dots = |x_k| = d$ and z is an additional tuple of free variables, we let $R_\varphi^*(n) := \max\{R_{\varphi(x_1, \dots, x_k; b)}(n) : b \in M^{|z|}\}$ (or ∞ if the maximum doesn't exist).

Remark 1.9. By Tarski's quantifier elimination, in $\mathcal{M} = (\mathbb{R}, <, +, \times, 0, 1)$, given a formula $\varphi(x; y) \in \mathcal{L}$, all sets of the form $\varphi(\mathbb{R}^{|x|}; b), b \in \mathbb{R}^{|y|}$ are semialgebraic of the description complexity $\leq t$ for some t depending only on φ . Conversely, the family of all semialgebraic subsets of $\mathbb{R}^{|x|}$ of the description complexity $\leq t$ is of the form $\{\varphi(x; b) : b \in \mathbb{R}^{|y|}\}$ for an appropriate choice of $\varphi(x; y) \in \mathcal{L}$. Hence $R_k^{d,t}$ from Definition 1.3 is given by R_φ^* for an appropriate φ in the case of the field of reals.

We will restrict to the case of *NIP structures* (see Section 2 for the definition; any structure which is not NIP codes arbitrary finite graphs in a definable way, hence bounds in Fact 1.2 are optimal outside of the NIP context). First we overview briefly the relevant results in the model-theoretic literature indicating the relevance of NIP and its subclasses for the problem at hand.

The infinitary version of the question was long known in model theory, under the name of the “*existence of indiscernibles*” (starting with the work of Morley in the stable case, and later work of Shelah and others in general NIP [24, 34–36]).

The question of obtaining explicit bounds for $R_\varphi(n)$ under some model-theoretic tameness assumptions on \mathcal{M} was first considered, it appears, in [14], where some quantitative improvements in the stable and NIP cases were obtained. In the case of a stable formula φ , a polynomial upper bound was established by Malliaris and Shelah.

Fact 1.10. [27] *Let $\varphi(x_1, \dots, x_k; z)$ be a formula in a stable structure \mathcal{M} (or just assume that φ is a stable formula, relatively to an arbitrary partition of its variables). Then there is some $c = c(\varphi)$ such that $R_\varphi^*(n) \leq n^c$ for all sufficiently large n .*

See also [8] for a different proof using the “non-standard” method. Fact 1.4 was generalized to *o-minimal* structures (with some additional topological assumptions) in [4], and to arbitrary distal structures in the following theorem.

Fact 1.11. [9] *Let \mathcal{M} be a distal structure. Then for any formula $\varphi(x_1, x_2; z)$, with $|x_1| = |x_2|$ arbitrary, there is some $c = c(\varphi)$ such that $R_\varphi^*(n) \leq n^c$ for all sufficiently large n .*

In this paper, we continue investigating the bounds for the functions $R_\varphi(n)$ and $R_\varphi^*(n)$ in various NIP structures. First, we consider an analog of the Bukh-Matousek theorem (Fact 1.7) in *o-minimal* structures. Recall that a structure $\mathcal{M} = (M, <, \dots)$ is *o-minimal* if every definable subset of M is a finite union of singletons and intervals (with endpoints in $M \cup \{\pm\infty\}$). From this assumption one obtains cell decomposition and other geometric information for definable subsets of M^n , for all n . The theory of *o-minimal* structures is rather well-developed and has applications in other branches of mathematics (we refer to [38] for a detailed treatment of *o-minimality*, or to [33, Section 3] and references there for a quick introduction). Examples of *o-minimal* structures include $\mathbb{R} = (\mathbb{R}, +, \times)$, $\mathbb{R}_{\exp} = (\mathbb{R}, +, \times, e^x)$,

$\mathbb{R}_{\text{an}} = \left(\mathbb{R}, +, \times, f \upharpoonright_{[0,1]^k} \right)$ for f ranging over all functions real-analytic on some neighborhood of $[0, 1]^k$, or the combination of both $\mathbb{R}_{\text{an,exp}}$. An o -minimal structure \mathcal{M} is *polynomially bounded* if for every definable function f , there exists $N \in \mathbb{N}$ such that $|f(x)| \leq x^N$ for all sufficiently large positive x . So for example \mathbb{R} and \mathbb{R}_{an} are polynomially bounded, but \mathbb{R}_{exp} is not. In Section 3 we generalize Fact 1.7 to arbitrary polynomially bounded o -minimal expansions of the field of reals \mathbb{R} .

Theorem 1.12. *Let \mathcal{M} be a polynomially bounded o -minimal expansion of \mathbb{R} . Then for every $k \in \mathbb{N}$ and every formula $\varphi(x_1, \dots, x_k; z)$ with x_1, \dots, x_k singletons, i.e. $|x_i| = 1$, there is some $c = c(\varphi)$ such that $R_\varphi^*(n) \leq 2^{2^{cn}}$ for all sufficiently large n .*

In particular this implies that in the semialgebraic case (Fact 1.7) the constant c only depends on the description complexity of the relation, and not on the magnitude of the parameters, which doesn't seem to have been noticed before. Our argument combines uniform definability of types over finite sets in NIP structures (see Definition 2.4) and a combinatorial lemma from [6]. On the other hand, in Section 4 we show that no analog of Theorem 1.12 can hold in \mathbb{R}_{exp} .

Theorem 1.13. *For every $k \geq 3$ there are relations $E_k(x_1, \dots, x_k)$ definable in \mathbb{R}_{exp} with $|x_i| = 1$, constants $C_k > 0$ and $n_k \in \mathbb{N}$ such that for each $n > n_k$ there is a sequence \tilde{a}_n in \mathbb{R} of length n that doesn't contain an E_k -indiscernible subsequence of length greater than $C_k \log \log \dots \log n$, with $k - 2$ iterations of \log (in this paper, \log always means logarithm with base 2).*

By a theorem of Miller [29], if an o -minimal expansion of the field of real numbers is not polynomially bounded, then exponentiation is definable in it. Hence the assumption in Theorem 1.12 is the most general possible among o -minimal structures.

In Section 5 we prove an analog of Fact 1.7 in the field of the p -adics \mathbb{Q}_p , for every prime p (and more generally, for p -minimal structures satisfying some additional technical assumptions).

Theorem 1.14. *For a prime p , let $\mathcal{M} = (\mathbb{Q}_p, +, \times, 0, 1)$ be the field of p -adics. Then for every $k \in \mathbb{N}$ and every formula $\varphi(x_1, \dots, x_k; z)$, with x_1, \dots, x_k singletons, there is some $c = c(\varphi)$ such that $R_\varphi^*(n) \leq 2^{2^{cn}}$ for all sufficiently large n .*

In Section 6 we consider the growth of $R_\varphi^*(n)$ in NIP structures for definable relations of higher arity. Generalizing Fact 1.5, we show a definable stepping down lemma for NIP structures which implies the following.

Theorem 1.15. *Let \mathcal{M} be an NIP structure, and assume that for all binary formulas $\varphi(x_1, x_2; z) \in \mathcal{L}$ we have $R_\varphi^*(n) \leq n^c$ for some $c = c(\varphi)$ and all n large enough. Then for all $k \geq 3$ and all $\varphi(x_1, \dots, x_k; z)$ we have $R_\varphi^*(n) \leq \text{twr}_{k-1}(n^c)$ for some $c = c(\varphi)$ and all n large enough.*

The assumption of Theorem 1.15 is satisfied in distal structures by Fact 1.11, and in stable structures by Fact 1.10. We conjecture that it also holds in arbitrary NIP structures, which is equivalent to saying that the Erdős-Hajnal conjecture holds for all graphs definable in NIP structures. We refer to [9, 10] for a discussion.

2. PRELIMINARIES ON NIP

Vapnik–Chervonenkis dimension, or VC-dimension, is an important notion in combinatorics and statistical learning theory (see e.g. [28] for an exposition). Let

X be a set, finite or infinite, and let \mathcal{F} be a family of subsets of X . Given $A \subseteq X$, we say that it is *shattered* by \mathcal{F} if for every $A' \subseteq A$ there is some $S \in \mathcal{F}$ such that $A \cap S = A'$. A family \mathcal{F} is a *VC-class* if there is some $n < \omega$ such that no subset of X of size n is shattered by \mathcal{F} . In this case *the VC-dimension of \mathcal{F}* , that we will denote by $VC(\mathcal{F})$, is the smallest integer n such that no subset of X of size $n + 1$ is shattered by \mathcal{F} . For a set $B \subseteq X$, let $\mathcal{F} \cap B = \{A \cap B : A \in \mathcal{F}\}$ and let $\pi_{\mathcal{F}}(n) = \max \{|\mathcal{F} \cap B| : B \subseteq X, |B| = n\}$.

Fact 2.1 (Sauer-Shelah lemma). *If $VC(\mathcal{F}) \leq d$ then for $n \geq d$ we have $\pi_{\mathcal{F}}(n) \leq \sum_{i \leq d} \binom{n}{i} = O(n^d)$.*

An important class of NIP theories was introduced by Shelah in his work on the classification program [34]. It has attracted a lot of attention recently, both from the point of view of pure model theory and due to some applications in algebra and geometry (see e.g. [1, 37] for an introduction to the area). Examples of NIP structures are given by arbitrary stable structures, ω -minimal structures, the field of p -adics for every prime p (along with its analytic expansion), as well as algebraically closed valued fields. As was observed in [25], the original definition of NIP is equivalent to the following one (see [3] for a more detailed account).

Definition 2.2. Let T be a complete theory and $\varphi(x, y)$ a formula in T , where x, y are tuples of variables, possibly of different length. We say that *the formula $\varphi(x, y)$ is NIP* if there is a model \mathcal{M} of T such that the family of sets $\{\varphi(M, a) : a \in M^{|y|}\}$ is a VC-class (as usual, $\varphi(M, a) = \{b \in M^{|x|} : \mathcal{M} \models \varphi(b, a)\}$). In this case we define the *VC-dimension of $\varphi(x, y)$* to be the VC-dimension of this class. (It is easy to see that by elementarily equivalence the above does not depend on the model \mathcal{M} of T .) A *theory T is NIP* if all formulas in T are NIP, and a structure \mathcal{M} is NIP if its complete theory $Th(\mathcal{M})$ is NIP. That is, a structure \mathcal{M} is NIP if for every formula $\varphi(x, y)$ the family of φ -definable sets $\mathcal{F}_{\varphi} = \{\varphi(M, a) : a \in M^{|y|}\}$ is a VC-class.

By a *partitioned* set of formulas $\Delta(x, y)$, where x and y are two groups of variables, we mean a set of formulas all of which are of the form $\varphi(x, y) \in \mathcal{L}$, i.e. have the same free variables partitioned into the same two groups. Given a (partitioned) set of formulas $\Delta(x, y)$ and a set $B \subseteq M^{|y|}$, we say that $\pi(x)$ is a *Δ -type over B* if $\pi(x) \subseteq \bigcup_{\varphi(x, y) \in \Delta, b \in B} \{\varphi(x, b), \neg \varphi(x, b)\}$ and there is some $\mathcal{N} \succeq \mathcal{M}$ and some $a \in N^{|x|}$ satisfying simultaneously all formulas from $\pi(x)$. By a *complete Δ -type over B* we mean a maximal Δ -type over B . We will denote by $S_{\Delta}(B)$ the collection of all complete Δ -types over B . If Δ consists of a single formula $\varphi(x, y)$, we simply say φ -type and write $S_{\varphi}(B)$, etc. In view of the remarks above, the following is an immediate corollary of the Sauer-Shelah lemma.

Fact 2.3. *A structure \mathcal{M} is NIP if and only if for any finite set of formulas $\Delta(x, y)$ there is some $d \in \mathbb{N}$ such that $|S_{\Delta}(B)| = O(|B|^d)$ for any finite $B \subseteq M^{|y|}$.*

This result can be strengthened. The following definition is from [3, 20].

Definition 2.4. (1) Given a $\varphi(x, y)$ -type $q \in S_{\varphi}(B)$, $B \subseteq M^{|y|}$, an $L(M)$ -formula $d\varphi(y)$ is said to *define q* if for all $b \in B$ we have

$$\varphi(x, b) \in q \iff \mathcal{M} \models d\varphi(b).$$

- (2) We say that $\varphi(x, y)$ -types are *uniformly definable over finite sets*, with d parameters, if there is a finite set of \mathcal{L} -formulas $\Delta = (d\varphi_i(y; y_1, \dots, y_d) : i < k)$ such that for every finite set $B \subseteq M^{|y|}$ and every $q \in S_\varphi(B)$ there are some $b_1, \dots, b_d \in B$ and some $i < k$ such that $d\varphi_i(y; b_1, \dots, b_d)$ defines q . We call the set Δ a *uniform definition* for φ -types over finite sets, with d parameters.
- (3) We say that T satisfies the *Uniform Definability of Types over Finite Sets*, or *UDTFS*, if for some (equivalently, any) $\mathcal{M} \models T$, φ -types are uniformly definable over finite sets for all formulas $\varphi \in L$.

Fact 2.5. [7] *Every NIP theory satisfies UDTFS.*

This result can be viewed as a model-theoretic version of the Warmuth conjecture on the existence of compression schemes for VC-families, which was later established in [31]. Special cases of Fact 2.5 were proved earlier for stable [34], ω -minimal [23], and dp-minimal [20] structures. Note that this implies Fact 2.3 since under UDTFS, for every finite set of formula Δ , every Δ -type over a finite set B is determined by fixing a definition for each $\varphi \in \Delta$ with parameters from B , of which there are only polynomially many choices. Explicit bounds on the number of parameters needed are given in [3] for some cases considered in this article.

- Fact 2.6.** (1) [3, Section 6.1] *Let \mathcal{M} be a (weakly or quasi) ω -minimal structure. Then $\varphi(x, y)$ -types are uniformly definable over finite sets using $|x|$ parameters, for all formulas $\varphi \in L$. In particular this applies to the Presburger arithmetic $(\mathbb{Z}, +, <)$.*
- (2) [3, Section 7.2] *Let \mathcal{M} be the field of p -adics. Then $\varphi(x, y)$ -types are uniformly definable over finite sets using $2|x|$ parameters, for all formulas $\varphi \in L$.*

Finally, we recall global invariant types and their products. Let $\mathbb{M} \succ M$ be a saturated elementary extension. We call complete types in $S_x(\mathbb{M})$ *global*, and we say that a global type p is *M -invariant* if it is $\text{Aut}(\mathbb{M}/M)$ -invariant (meaning that for every automorphism σ of \mathbb{M} fixing M pointwise, for every $L(\mathbb{M})$ -formula $\varphi(x, a)$ we have $\varphi(x, a) \in p \iff \varphi(x, \sigma(a)) \in p$).

Fact 2.7. (see e.g. [22, Section 2] or [37]) *Let p be a global M -invariant type. Let the sequence $(c_i : i \in \mathbb{N})$ in \mathbb{M} be such that $c_i \models p|_{M c_{<i}}$ (such a sequence is called a Morley sequence in p over M). Then the sequence $(c_i : i \in \mathbb{N})$ is indiscernible over M and $\text{tp}((c_i : i \in \mathbb{N})/M)$ does not depend on the choice of (c_i) . Call this type $p^{(\omega)}|_M$, and let $p^{(n)}|_M := \text{tp}(c_1, \dots, c_n/M)$.*

3. BUKH-MATOUSEK THEOREM IN POLYNOMIALLY BOUNDED ω -MINIMAL EXPANSIONS OF \mathbb{R}

First we prove a general lemma about NIP structures, which is a finitary version of Shelah's “shrinking of indiscernibles”.

Lemma 3.1. *Let \mathcal{M} be an NIP structure, and let $\varphi(x_1, \dots, x_n; y)$ be an arbitrary formula, with $|x_1| = \dots = |x_n| = d$. Then there are some $k, l \in \mathbb{N}$ and a finite set of formulas Δ in the variables x_1, \dots, x_l with $|x_i| = d$ such that for any finite Δ -indiscernible sequence $(a_i)_{i < N}$ in M^d and any $b \in M^{|y|}$ there are $0 = j_0 < j_1 < \dots < j_{k'} = N - 1$ with $k' \leq k$ such that for every $s \in \{0, \dots, k' - 1\}$ the sequence $(a_i : j_s < i < j_{s+1})$ is $\varphi(x_1, \dots, x_n, b)$ -indiscernible.*

In particular, for any $b \in M^{|y|}$, any finite Δ -indiscernible sequence of elements in M^d of length N contains a $\varphi(x_1, \dots, x_n, b)$ -indiscernible subsequence of length at least $\frac{N-(k+1)}{k}$.

Proof. To simplify the notation we assume $d = 1$.

By UDTFs (Fact 2.5) applied to the formula $\varphi^{\text{op}}(y; x_1, \dots, x_n) := \varphi(x_1, \dots, x_n; y)$, there is a finite set of formulas $\Delta(x_1, \dots, x_n; \bar{x}_1, \dots, \bar{x}_m)$ with $|\bar{x}_i| = n$ such that for any finite set $A \subseteq M$ and $b \in M^{|y|}$ the φ^{op} -type of b over A^n is definable by an instance of some $\psi \in \Delta$ with parameters from A^n . That is, there are some $\bar{c}_1, \dots, \bar{c}_m \in A^n$, such that for all $a_1, \dots, a_n \in A$ we have $\models \varphi(a_1, \dots, a_n; b)$ if and only if $\models \psi(a_1, \dots, a_n; \bar{c}_1, \dots, \bar{c}_m)$.

Writing each n -tuple \bar{x}_i , $i = 1, \dots, m$ as n single variables in every $\psi \in \Delta$, we can view Δ as a finite set of formulas in the variables x_1, \dots, x_l , where $l = n + mn$.

Let $(a_i)_{i < N}$ be a finite Δ -indiscernible sequence, $b \in M^{|y|}$, and $A = \{a_i : i < N\}$. We choose $\psi \in \Delta$ and $c_{n+1}, \dots, c_l \in A$ such that for all $c_1, \dots, c_n \in A$ we have $\mathcal{M} \models \varphi(c_1, \dots, c_n; b)$ if and only if $\mathcal{M} \models \psi(c_1, \dots, c_n, c_{n+1}, \dots, c_l)$.

We choose $0 = j_0 < j_1 < \dots < j_{k'} = N - 1$ with $k' \leq (l - n) + 2 = mn + 2$ so that $\{a_{j_s} : s = 0, \dots, k'\} = \{c_i : i = n + 1, \dots, l\} \cup \{a_0, a_{N-1}\}$.

Since $(a_i)_{i < N}$ is ψ -indiscernible, it follows that for any $0 \leq i_1 < \dots < i_n < N$ the truth value of $\psi(a_{i_1}, \dots, a_{i_n}; c_{n+1}, \dots, c_l)$, and so of $\varphi(a_{i_1}, \dots, a_{i_n}; b)$, is determined by the quantifier-free order type of (i_1, \dots, i_n) over $\{j_s : s = 0, \dots, k'\}$. The conclusion of the lemma follows. \square

From now on we work in a polynomially bounded \mathcal{o} -minimal expansion $\mathcal{R} = \langle \mathbb{R}, <, \dots \rangle$ of the field of real numbers. Let $T = Th(\mathcal{R})$ and let $\mathbb{M} \succ \mathcal{R}$ be a big saturated model.

As T has Skolem functions (see e.g. [38]), it follows that for all $M \prec \mathbb{M}$ and $\bar{a} \in \mathbb{M}^n$, the set

$$M\langle \bar{a} \rangle = \{f(\bar{a}) : f(x) \text{ is an } M\text{-definable function}\}$$

is an elementary substructure of \mathbb{M} .

Let $\tilde{p}(x) \in S_1(\mathbb{M})$ be the global type of “ $+\infty$ ”, i.e. \tilde{p} is the unique complete global type such that $\tilde{p} \vdash x > m$ for every $m \in \mathbb{M}$ (uniqueness is by \mathcal{o} -minimality). It is invariant over \emptyset (as the set of formulas $\{m < x : m \in \mathbb{M}\}$ is clearly $\text{Aut}(\mathbb{M}/\emptyset)$ -invariant).

The following fact is obvious.

Fact 3.2. *For every $M \prec \mathbb{M}$, an element $\alpha \in \mathbb{M}$ realizes $\tilde{p}(x)|_M$ if and only if $\alpha > m$ for every $m \in M$.*

Since polynomial boundedness is preserved under elementary equivalence (see [30, Theorems A and B]) we have the following fact.

Fact 3.3. *If $M \prec \mathbb{M}$ and $\alpha \models \tilde{p}|_M$, then the set $\{\alpha^n : n \in \mathbb{N}\}$ is cofinal in $M\langle \alpha \rangle$, i.e. for every $m \in M\langle \alpha \rangle$ there is some $n \in \mathbb{N}$ such that $m < \alpha^n$.*

Lemma 3.4. *Let $M \prec \mathbb{M}$ and $\alpha_1, \dots, \alpha_n \in \mathbb{M}$. Then $(\alpha_1, \dots, \alpha_n)$ realizes $\tilde{p}^{(n)}|_M$ if and only if $\alpha_1 > m$ for all $m \in M$ and $\alpha_{i+1} > \alpha_i^k$ for all $k \in \mathbb{N}$ and $i = 1, \dots, n-1$.*

Proof. Let $M_0 = M$, and for $i = 1, \dots, n-1$ let $M_i = M_{i-1}\langle \alpha_i \rangle$.

Obviously for any $A \subset \mathbb{M}$ an element $\alpha \in \mathbb{M}$ realizes $\tilde{p}|_A$ if and only if it realizes $\tilde{p}|_{\text{dcl}(A)}$. Thus $(\alpha_1, \dots, \alpha_n)$ realizes $\tilde{p}^{(n)}|_M$ if and only if α_{i+1} realizes $\tilde{p}|_{M_i}$ for $i = 0, \dots, n-1$, and Lemma follows from facts 3.2 and 3.3 \square

In view of the above lemma, we define “finitary” approximations to a realization of $\tilde{p}^{(n)}|_{\mathcal{R}}$.

Definition 3.5 (Definition 2.1 [6]). Let $R > 2$ be a real number. A sequence $\vec{a} = (a_1, \dots, a_n)$ in \mathcal{R} is called *R-growing* if $a_1 \geq R$ and $a_{i+1} \geq a_i^R$ for $i = 1, \dots, n-1$.

Notice that any subsequence of an *R-growing* sequence is *R-growing* as well.

Lemma 3.6. *For any finite set of formulas $\Delta(x_1, \dots, x_l)$ with parameters from \mathbb{R} there is some $R \in \mathbb{R}$ such that any *R-growing* sequence of elements $a_i \in \mathbb{R}, i = 1, \dots, N$ is Δ -indiscernible.*

Proof. Consider the partial type

$$\Sigma(x_1, \dots, x_{2l}) = \left\{ x_1 > n \wedge \bigwedge_{i=1}^{2l-1} (x_{i+1} > x_i^n) : n \in \mathbb{N} \right\}.$$

By Fact 2.7, for any $N \in \mathbb{N}$, if $(a_1, \dots, a_N) \models \tilde{p}^{(N)}|_M$, then the sequence (a_1, \dots, a_N) is indiscernible. Together with Lemma 3.4 this implies that

$$\Sigma(x_1, \dots, x_{2l}) \vdash \psi(x_1, \dots, x_l) \leftrightarrow \psi(x_{i_1}, \dots, x_{i_l})$$

for any $1 \leq i_1 < i_2 < \dots < i_l \leq 2l$ and $\psi \in \Delta$. By compactness, this holds with Σ replaced by some finite subset Σ_0 . But then, if a_1, \dots, a_N is an *R-growing* sequence and R is larger than the largest n appearing in Σ_0 , then every increasing $2l$ -tuple from a_1, \dots, a_N satisfies Σ_0 , hence a_1, \dots, a_N is $\Delta(x_1, \dots, x_l)$ -indiscernible. \square

Combining Lemma 3.6 with Lemma 3.1 we can allow additional parameters in Δ .

Corollary 3.7. *For any finite set of formulas $\Delta(x_1, \dots, x_l; y)$ with parameters from \mathbb{R} there is some $R \in \mathbb{R}$ and $m \in \mathbb{N}$ such that for any *R-growing* sequence of elements $\vec{a} = (a_i : i = 1, \dots, N)$ in \mathbb{R} with N large enough and any $b \in \mathbb{R}^{|y|}$, \vec{a} contains a $\Delta(x_1, \dots, x_l; b)$ -indiscernible subsequence of length $\frac{N}{m}$.*

Proof. For every $\varphi(x_1, \dots, x_l; y) \in \Delta$, let $k_\varphi \in \mathbb{N}$ and Δ_φ a finite set of formulas be as given by Lemma 3.1 for φ , and let $\Delta' = \bigcup_{\varphi \in \Delta} \Delta_\varphi$ and $k = \max\{k_\varphi : \varphi \in \Delta\}$. Now by Lemma 3.6 there is some R such that every *R-growing* sequence $\vec{a} = (a_1, \dots, a_N)$ of elements from \mathbb{R} is Δ' -indiscernible. By Lemma 3.1, for any $b \in \mathbb{R}^{|y|}$ we can find an interval $[i^0, i^1]$ in $[1, N]$ of length at least $\frac{N - (k|\Delta| - 2)}{k|\Delta|}$ such that the sequence $(a_i : i^0 \leq i \leq i^1)$ is $\Delta(x_1, \dots, x_l; c)$ -indiscernible. We can take $m = 2k|\Delta|$. \square

Finally, the following combinatorial lemma is from [6] (namely, Proposition 2.4 combined with Definition 2.3 there).

Fact 3.8. *For every n and $R \geq R_0$, where R_0 is a certain absolute constant, there exists $N \leq 2^{R^{2n}}$ such that for any sequence \vec{a} of length N there is an *R-growing* sequence \vec{b} of length n and $A, B \in \mathbb{R}$ such that one of the following sequences is a subsequence of \vec{a} .*

- (1) $A + Bb_i, i = 1, \dots, n.$
- (2) $A + \frac{B}{b_i}, i = 1, \dots, n.$
- (3) $A + Bb_i, i = n, \dots, 1.$
- (4) $A + \frac{B}{b_i}, i = n, \dots, 1.$

(Note: the order in (3) and (4) is reversed.)

We are ready to prove the main result of the section, generalizing [6, Proposition 1.6].

Theorem 3.9. *Let \mathcal{R} be a polynomially bounded o-minimal expansion of the real field. Then for any formula $\varphi(x_1, \dots, x_r; z)$ with parameters from \mathbb{R} , with all x_i singletons, there is a constant $C = C(\varphi)$ such that*

$$R_\varphi^*(n) \leq 2^{2^{Cn}},$$

for all sufficiently large n .

Proof. Let $\Delta(x_1, \dots, x_r; y_1, y_2, z)$ consist of the formulas

$$\varphi_1(x_1, \dots, x_r; y_1, y_2, z) = \varphi(y_1 + y_2 x_1, \dots, y_1 + y_2 x_r; z),$$

$$\varphi_2(x_1, \dots, x_r; y_1, y_2, z) = \varphi(y_1 + \frac{y_2}{x_1}, \dots, y_1 + \frac{y_2}{x_r}; z),$$

$$\varphi_3(x_1, \dots, x_r; y_1, y_2, z) = \varphi(y_1 + y_2 x_r, \dots, y_1 + y_2 x_1; z),$$

$$\varphi_4(x_1, \dots, x_r; y_1, y_2, z) = \varphi(y_1 + \frac{y_2}{x_r}, \dots, y_1 + \frac{y_2}{x_1}; z),$$

and let R and m be as given by Corollary 3.7 for Δ . Now assume that \vec{a} is an arbitrary sequence of singletons of length $N = 2^{R^{2mn}}$ (which is bounded by $2^{2^{Cn}}$ for an appropriate constant C depending just on m, R), and $e \in \mathbb{R}^{|z|}$.

By Fact 3.8, there is some R -growing sequence $\vec{b} = (b_i : 1 \leq i \leq kn)$ and some $A, B \in \mathbb{R}$ such that one of the corresponding sequences given by (1)–(4) in Fact 3.8 is a subsequence of \vec{a} . By Corollary 3.7, \vec{b} contains a $\Delta(x_1, \dots, x_r; A, B, e)$ -indiscernible subsequence of length n . But by the choice of Δ , the corresponding subsequence of \vec{a} must be $\varphi(x_1, \dots, x_r; e)$ -indiscernible. \square

4. COUNTEREXAMPLE IN \mathbb{R}_{exp}

4.1. Preliminaries. We work in the structure \mathbb{R}_{exp} in the language $\mathcal{L} = (<, +, \times, 0, 1, \exp(x))$, i.e the expansion of the field of reals with the exponential function. It is well-known to be o-minimal [39].

As in Definition 1.8, for an $\mathcal{L}(\mathcal{M})$ -formula $\varphi(x_1, \dots, x_k)$ with dimension $|x_1| = \dots = |x_k| = d$, for an integer $n \geq k$ we denote by $R_\varphi(n)$ the smallest integer N such that any sequence of elements in \mathbb{R}^d of length N contains a φ -indiscernible subsequence of length n .

Instead of tower notations we use iterated log and exp. By induction on n we define $e_n(x)$ and $l_n(x)$ as

$$e_0(x) = x, e_{n+1}(x) = 2^{e_n(x)}; \text{ and } l_0(x) = x, l_{n+1}(x) = \log(l_n(x)),$$

where by log we always mean \log_2 . Obviously $l_n(x)$ is defined for large enough x and it is the compositional inverse of $e_n(x)$.

Our goal is to prove the following theorem.

Theorem 4.1. *For every $k \geq 3$ there is a relation $E_k(x_1, \dots, x_k)$ definable in \mathbb{R}_{exp} with $|x_1| = \dots = |x_k| = 1$ and $c_k > 0$ such that $R_{E_k}(n) \geq e_{k-2}(c_k n)$ for all sufficiently large n .*

The proof of the above theorem closely follows the proof of Theorem 1.2 in [11] (see also Theorem 1.3 in [13]), by showing that in the structure \mathbb{R}_{exp} the stepping-up approach of Erdős and Hajnal can be implemented definably without increasing the dimension. First we discuss some preliminaries.

4.2. Robustness. We will use the notion of robustness from [13] (that was originally called “depth” in [11]).

Definition 4.2. Let $\varphi(x_1, \dots, x_k)$ be an \mathcal{L} -formula and $\vec{a} = (a_1, \dots, a_n)$ be a sequence of real numbers. We say that φ is *robust* on \vec{a} if there is $\varepsilon > 0$ such that, for all $1 \leq i_1 < \dots < i_k \leq n$ and all real numbers a'_1, \dots, a'_k with $|a_{i_j} - a'_j| < \varepsilon$ for each $j = 1, \dots, k$, we have

$$\models \varphi(a'_1, \dots, a'_k) \leftrightarrow \varphi(a_{i_1}, \dots, a_{i_k}).$$

4.3. \log_T -transformations.

Definition 4.3. Let $\varphi(x_1, \dots, x_r)$ be an \mathcal{L} -formula. Let $T > 0$ be a real number. For a formula $\psi(y_1, \dots, y_s)$ we say that ψ is a \log_T -transformation of φ if it is obtained from φ by replacing **every** free variable x_i in φ by an expression of the form $\log_T(u_i - v_i)$ with $u_i, v_i \in \{y_1, \dots, y_s\}$.

Definition 4.4. We say that an \mathcal{L} -formula $\varphi(x_1, \dots, x_r)$ is an *rd-formula* if it depends only on the ratios of differences of its variables, i.e. it is equivalent to a formula of the form

$$\psi\left(\frac{x_{i_1} - x_{j_1}}{x_{p_1} - x_{q_1}}, \dots, \frac{x_{i_s} - x_{j_s}}{x_{p_s} - x_{q_s}}\right)$$

for some $\psi(y_1, \dots, y_s) \in \mathcal{L}$, where $i_t, j_t, p_t, q_t \in \{1, \dots, r\}$ for all $t = 1, \dots, s$ (and there are no other free variables in ψ).

Claim 4.5. *Let $T > 0$. If $\psi(y_1, \dots, y_s)$ is a \log_T -transformation of an rd-formula $\varphi(x_1, \dots, x_r)$, then ψ is also an rd-formula, and it is also a \log_2 -transformation of φ (with the same choices of $u_i, v_i \in \{y_1, \dots, y_s\}$ for $1 \leq i \leq r$).*

Proof. Applying \log_T -substitution to an expression of the form $\frac{x_i - x_j}{x_p - x_q}$ we obtain the expression

$$\frac{\log_T(u_i - v_i) - \log_T(u_j - v_j)}{\log_T(u_p - v_p) - \log_T(u_q - v_q)},$$

which is equivalent to

$$\frac{\log_T \frac{u_i - v_i}{u_j - v_j}}{\log_T \frac{u_p - v_p}{u_q - v_q}}.$$

Since the ratio of two logarithms does not depend on the base, it is also equivalent to

$$\frac{\log \frac{u_i - v_i}{u_j - v_j}}{\log \frac{u_p - v_p}{u_q - v_q}}.$$

Thus, after applying \log_T -substitutions, we obtain an rd-formula that is also a \log_2 -substitution. \square

4.4. Proof of the theorem 4.1. For a formula $\varphi(x_1, \dots, x_k) \in \mathcal{L}$ with $|x_1| = \dots = |x_k| = d$ and an integer n we will denote by $R_\varphi^+(n)$ the smallest integer N such that any increasing sequence $a_1 < \dots < a_N$ contains a φ -indiscernible subsequence of length n .

Obviously for any formula $\varphi(x_1, \dots, x_k)$ with $|x_i| = 1$ we have $R_\varphi^+(n) \leq R_\varphi(n)$.

Thus Theorem 4.1 follows from the following refined version.

Theorem 4.6. *For every $k \geq 3$ there are an rd-formula $E_k(x_1, \dots, x_k) \in \mathcal{L}$ with $|x_1| = \dots = |x_k| = 1$ and a constant $C_k > 1$ such that, for all real $0 < c < 1$ and for all large enough $n \in \mathbb{N}$, there is an increasing sequence of natural numbers \vec{a}^n of length at least $e_{k-2}(cn)$ such that E_k is robust on \vec{a}^n , and \vec{a}^n does not contain an E_k -indiscernible subsequence of length $C_k n$.*

Proof of 4.6 \Rightarrow 4.1. Fix $0 < c < 1$ and set $c_k = \frac{c}{C_k}$, for each $k \geq 3$, where C_k is the constant given by 4.6. We then have, for the rd-formula $E_k(x_1, \dots, x_k)$ given to us by 4.6 and for all large enough n , an increasing sequence of natural numbers \vec{a}^n of length at least $e_{k-2}(c_k n)$ such that \vec{a}^n does not contain an E_k -indiscernible subsequence of length n . Thus $R_\varphi^+(n) \geq e_{k-2}(c_k n)$, and hence $R_\varphi(n) \geq e_{k-2}(c_k n)$ by the preceding remark. \square

Remark 4.7. To prove Theorem 4.6 we only need to construct formulas $E_k(x_1, \dots, x_k)$ whose truth values are well-defined only on increasing sequences of real numbers $r_1 < \dots < r_k$. (The formula $\log(x_2 - x_1) > \log(x_3 - x_2)$ is an example of a formula that we will use often.)

We proceed by induction on k .

4.5. The base case $k = 3$. For the following claim see [11, Section 3.1].

Claim 4.8. *Let $E_3(x_1, x_2, x_3)$ be the formula $x_1 + x_3 - 2x_2 \geq 0$. Then for any $n \geq 1$ the sequence $1, 2, 3, \dots, 2^n$ does not contain an E_3 -indiscernible subsequence of length $n + 2$.*

It is not hard to see that E_3 is equivalent to an rd-formula. Indeed we can rewrite E_3 as $x_3 - x_2 > x_2 - x_1$, which on increasing sequences is equivalent to $\frac{x_3 - x_2}{x_2 - x_1} > 1$.

We also need E_3 to be robust on \vec{a}^n . It is not hard to see that E_3 is not robust on the sequence $1, 2, \dots, 2^n$, since $1 + 3 - 2 \cdot 2 = 0$ and the truth of E_3 can change even if we perturb the first 3 elements of the sequence by arbitrarily small positive amounts. It is however also easy to see that E_3 is robust on any sequence that does not contain any terms $a < b < c$ with $a + c - 2b = 0$, i.e. it is robust on any sequence that does not contain a non-trivial 3-term arithmetic progression. To get such a sequence we use Behrend's Theorem (see [5]).

Theorem 4.9 (Behrend's Theorem). *There is a constant $D > 0$ such that for all natural numbers m there exists a set $X \subseteq \{1, \dots, m\}$ with $|X| \geq \frac{m}{2^{D\sqrt{\log m}}}$ not containing any non-trivial 3-term arithmetic progressions.*

For any $0 < c < 1$ and for all n large enough, $2^{n-D\sqrt{\log 2^n}} > 2^{cn}$. Therefore, for all large enough n , the sequence $1, 2, \dots, 2^n$ contains a subsequence of length 2^{cn} that does not contain a non-trivial 3-term arithmetic progression.

This finishes the case $k = 3$, taking $C_3 = 2^\eta$, for any $\eta > 0$ (as then $C_3 n \geq n + 2$ for all large enough n).

4.6. Inductive Step. Assume we have an rd-formula $E_k(x_1, \dots, x_k)$ as in Theorem 4.6. To complete the inductive step it is enough to construct an rd-formula $E_{k+1}(x_1, \dots, x_{k+1})$ satisfying the following:

Let \vec{a} be an increasing sequence of natural numbers of length N such that E_k is robust on \vec{a} , and \vec{a} does not contain an E_k -indiscernible subsequence of length n . Then there is an increasing sequence of natural numbers \vec{b} of length 2^N such that E_{k+1} is robust on \vec{b} and \vec{b} does not contain an E_{k+1} -indiscernible subsequence of length $2n + k - 4$. (We are then done taking $C_k = 2^{k-3+\eta}$, for $\eta > 0$ as fixed in the base case.)

Let $\vec{a} = (a_1, \dots, a_N)$ be an increasing sequence of natural numbers such that E_k is robust on \vec{a} , and \vec{a} does not contain an E_k -indiscernible sequence of length n .

Let T be a very large integer, specified later.

Consider the set

$$B_T = \left\{ \sum_{i=1}^N \beta_i T^{a_i} : \beta_i \in \{0, 1\} \right\}.$$

Since T is large enough, any $b \in B_T$ can be written uniquely as $b = \sum_{i=1}^N b(i) T^{a_i}$ with $b(i) \in \{0, 1\}$. Obviously B_T has size 2^N and we construct the sequence \vec{b}_T by taking the increasing enumeration of B_T .

For $b \neq c \in B_T$ let $\Delta(b, c) = \max\{i : b(i) \neq c(i)\}$. It is easy to see that, when T is large enough, for $b \neq c \in B_T$ and $i = \Delta(b, c)$ we have $b < c \Leftrightarrow b(i) < c(i)$.

Finally for $b \neq c \in B_T$ let $\delta(b, c) = a_{\Delta(b, c)}$.

We will now construct the step-up relation $E_k^\uparrow(x_1, \dots, x_{k+1})$ (not definable in \mathbb{R}_{exp}) on increasing $(k+1)$ -tuples of elements of B_T (we don't care how it is defined on the other elements).

Let $b_1 < b_2 < \dots < b_{k+1}$ be elements of B_T and for $i = 1, \dots, k$ let $\delta_i = \delta(b_{i+1}, b_i)$. Notice that δ_i is an element of \vec{a} .

We define $E_k^\uparrow(b_1, \dots, b_{k+1})$ to be true if and only if

$$\begin{aligned} E_k(\delta_1, \dots, \delta_k) \quad & \text{and} \quad \delta_1 < \delta_2 < \dots < \delta_k, \\ & \text{or} \\ E_k(\delta_k, \dots, \delta_1) \quad & \text{and} \quad \delta_1 > \delta_2 > \dots > \delta_k, \\ & \text{or} \\ \delta_1 < \delta_2 \quad & \text{and} \quad \delta_2 > \delta_3. \end{aligned}$$

It follows from the Erdős-Hajnal argument (see [11, Lemma 3.1]) that \vec{b}_T does not contain an E_k^\uparrow -indiscernible sequence of length $2n + k - 4$.

4.6.1. Definability. Now, as in [11], for $b > c$ we define

$$\bar{\delta}_T(b, c) = \log_T(b - c).$$

It is not hard to see that for any fixed $\varepsilon > 0$, if T is large enough, then for all $b > c \in B_T$ we have $|\delta(b, c) - \log_T(b - c)| < \varepsilon$.

Since E_k is robust on \vec{a} , choosing a very large integer T and considering the relation $E_k^{\uparrow T}(x_1, \dots, x_{k+1})$ obtained from E_k^\uparrow by replacing δ_i by $\bar{\delta}_T(b_{i+1}, b_i)$, we obtain that for $b_1 < \dots < b_{k+1} \in B_T$ we have $E_k^{\uparrow T}(b_1, \dots, b_{k+1})$ if and only if $E_k^\uparrow(b_1, \dots, b_{k+1})$. Hence \vec{b}_T does not contain an $E_k^{\uparrow T}$ -indiscernible sequence of length $2n + k - 4$.

Notice that $E_k^{\uparrow T}$ is definable and for $b_1 < b_2 < \dots < b_{k+1}$ we have that $E_k^{\uparrow T}(b_1, \dots, b_{k+1})$ holds if and only if

$$\begin{array}{lll} E_k(\bar{\delta}_T(b_2, b_1), \dots, \bar{\delta}_T(b_{k+1}, b_k)) & \text{and} & \bigwedge_{i=1}^{k-1} \bar{\delta}_T(b_{i+1}, b_i) < \bar{\delta}_T(b_{i+2}, b_{i+1}) \\ & \text{or} & \\ E_k(\bar{\delta}_T(b_{k+1}, b_k), \dots, \bar{\delta}_T(b_2, b_1)) & \text{and} & \bigwedge_{i=1}^{k-1} \bar{\delta}_T(b_{i+1}, b_i) > \bar{\delta}_T(b_{i+2}, b_{i+1}) \\ & \text{or} & \\ \bar{\delta}_T(b_2, b_1) < \bar{\delta}_T(b_3, b_2) & \text{and} & \bar{\delta}_T(b_3, b_2) > \bar{\delta}_T(b_4, b_3). \end{array}$$

Claim 4.10. $E_k^{\uparrow T}$ is equivalent to an rd-formula and does not depend on T .

Proof. By definition, $E_k^{\uparrow T}$ is a Boolean combination of \log_T -transformations of E_k and formulas of the form $\log_T(y - x) > \log_T(u - v)$.

By Claim 4.5, a \log_T -transformation of an rd-formula is an rd-formula that does not depend on T , and we only need to check that $\log_T(y - x) > \log_T(u - v)$ is equivalent to an rd-formula that does not depend on T . Indeed, $\log_T(y - x) - \log_T(u - v) > 0$ is equivalent to $\frac{y-x}{u-v} > 1$, which is an rd-formula. \square

Using Claim 4.10, we define E_{k+1} to be $E_k^{\uparrow 2}$. We can write a more explicit definition of E_{k+1} . It is the disjunction of three formulas $\varphi_1 \vee \varphi_2 \vee \varphi_3$, where

$$\begin{aligned} \varphi_1 & \text{ is } E_k\left(\log(x_2 - x_1), \dots, \log(x_{k+1} - x_k)\right) \wedge \bigwedge_{i=1}^{k-1} \left(\frac{x_{i+1} - x_i}{x_{i+2} - x_{i+1}} < 1\right), \\ \varphi_2 & \text{ is } E_k\left(\log(x_{k+1} - x_k), \dots, \log(x_2 - x_1)\right) \wedge \bigwedge_{i=1}^{k-1} \left(\frac{x_{i+1} - x_i}{x_{i+2} - x_{i+1}} > 1\right), \end{aligned}$$

and

$$\varphi_3 \text{ is } \frac{x_2 - x_1}{x_3 - x_2} < 1 \wedge \frac{x_2 - x_1}{x_3 - x_2} > 1.$$

It remains to show that for large enough T the relation E_{k+1} is robust on \vec{b}_T .

4.6.2. Robustness. It is not hard to see that since E_k is robust on \vec{a} and \log is continuous, both $E_k(\log_T(x_2 - x_1), \dots, \log_T(x_{k+1} - x_k))$ and $E_k(\log_T(x_{k+1} - x_k), \dots, \log_T(x_2 - x_1))$ are robust on \vec{b}_T , and we only need to check that all of the formulas $x_{i+1} - x_i < x_{i+2} - x_{i+1}$ and $x_{i+1} - x_i > x_{i+2} - x_{i+1}$ are robust on \vec{b}_T , i.e. for $b < c < d$ in B_T we don't have $c - b = d - c$. It is easy to check that there are no such b, c, d in B_T .

5. BUKH-MATOUSEK THEOREM IN THE p -ADICS (AND p -MINIMAL STRUCTURES)

Let p be a prime number. In this section we give an analog of Theorem 3.9 for relations definable in the field of p -adic numbers \mathbb{Q}_p and some of its expansions.

Let \mathcal{L}_p be the Macintyre language for the p -adics [26], i.e. \mathcal{L}_p consists of

- (a) the language of rings: $0, 1, +, -, \cdot, ^{-1}$,
- (b) a unary predicate V ,
- (c) a unary predicate P_n for each $n \in \mathbb{N}$;

with the usual interpretations in \mathbb{Q}_p : $V(\mathbb{Q}_p) = \mathbb{Z}_p$, and $P_n(\mathbb{Q}_p) = \{x \in \mathbb{Q}_p : \exists y \, x = y^n\}$. We will denote by T_p the complete theory $Th(\mathbb{Q}_p)$. By a result of Macintyre (see [26]), the theory T_p eliminates quantifiers in the language \mathcal{L}_p .

Similarly to the o -minimal case, there is a notion of minimality for expansions of \mathbb{Q}_p . Namely, an expansion $\tilde{\mathbb{Q}}_p$ in a language $\mathcal{L} \supseteq \mathcal{L}_p$ is p -minimal if in every

model of $Th(\tilde{\mathbb{Q}}_p)$, every definable subset in one variable is quantifier-free definable just using the language \mathcal{L}_p [21].

From now on, we fix a p -minimal expansion $\tilde{\mathbb{Q}}_p$ whose complete theory will be denoted by T , and we also fix a large sufficiently saturated and homogeneous model \mathbb{M} of T . We are following the same strategy as in Section 3. First we isolate some sufficiently representative global invariant types (in the o -minimal case, working with a single type of “ $+\infty$ ” was sufficient).

Proposition 5.1. *Let $\mathcal{M} \prec \mathbb{M}$ be a small model and let $\alpha_1, \alpha_2 \in \mathbb{M}$ be singletons with $\alpha_1 \equiv_{\emptyset} \alpha_2$, and $v(\alpha_l) > v(m)$ for $l = 1, 2$ and every $m \in M$. Then $\alpha_1 \equiv_M \alpha_2$.*

Proof. Since α_1 and α_2 are singletons, by p -minimality, we need to show the following:

- (1) $p(\alpha_1) = 0$ if and only if $p(\alpha_2) = 0$ for any polynomial $p(x) \in M[x]$;
- (2) $\models V(p(\alpha_1)/q(\alpha_1))$ if and only if $\models V(p(\alpha_2)/q(\alpha_2))$, for any $p(x), q(x) \in M[x]$;
- (3) $\models P_n(p(\alpha_1)/q(\alpha_1))$ if and only if $\models P_n(p(\alpha_2)/q(\alpha_2))$, for any $n \geq 2$ and $p(x), q(x) \in M[x]$.

Now (1) is easy, since by the ultrametric inequality, both α_1 and α_2 are transcendental over M . And (2) is equivalent to:

$$v(p(\alpha_1)) \geq v(q(\alpha_1)) \text{ if and only if } v(p(\alpha_2)) \geq v(q(\alpha_2)),$$

for any $p(x), q(x) \in M[x]$. Let $p(x) = a_0 + a_1x + \dots + a_kx^k$ and $q(x) = b_0 + b_1x + \dots + b_sx^s$. Let i be minimal with $a_i \neq 0$ and j be minimal with $b_j \neq 0$. Then, for $l = 1, 2$ we have $v(p(\alpha_l)) = v(a_i\alpha_l^i) = v(a_i) + iv(\alpha_l)$, and $v(q(\alpha_l)) = v(b_j\alpha_l^j) = v(b_j) + jv(\alpha_l)$. Thus $v(p(\alpha_l)) \geq v(q(\alpha_l))$ if and only if $i > j$ or $i = j$ and $v(a_i) \geq v(b_j)$. The latter condition is independent of l .

Finally, we demonstrate (3). It is easy to see that $\models P_n(p(\alpha_l)/q(\alpha_l))$ if and only if $\models P_n(p(\alpha_l)q^{n-1}(\alpha_l))$. Thus we need to show that $\models P_n(p(\alpha_1))$ if and only if $\models P_n(p(\alpha_2))$, for any $p(x) \in M[x]$.

We will need the following fact that follows easily from Hensel’s lemma.

Fact 5.2. *If $\varepsilon \in \mathbb{M}$ satisfies $v(\varepsilon) > k$ for all $k \in \mathbb{N}$ then for any $n \in \mathbb{N}$ the element $1 + \varepsilon$ has n -th root.*

Let $p(x) = a_0 + a_1x + \dots + a_kx^k$ be a nonzero polynomial over M and choose minimal i such that $a_i \neq 0$. Then, for $l = 1, 2$ we have $p(\alpha_l) = a_i\alpha_l^i(1 + \varepsilon_l)$ with $v(\varepsilon_l) > \mathbb{N}$, and $p(\alpha_l)$ has n -th root if and only if $a_i\alpha_l^i$ had n -th root. We can find $b_i \in M$ and $c \in \mathbb{Z}$ such that $a_i = b_i c$. Hence $a_i\alpha_l^i$ has n -th root if and only if $c\alpha_l^i$ does. Since \mathbb{Z} is in the definable closure of \emptyset , we have $\alpha_1 \equiv_{\mathbb{Z}} \alpha_2$ and $P_n(c\alpha_1)$ if and only if $P_n(c\alpha_2)$. \square

We now make some further assumptions on T (all of which are of course satisfied by \mathbb{Q}_p).

Extra Assumptions:

- (a) T has definable Skolem functions.
- (b) Whenever $\mathcal{M} \prec \mathbb{M}$ is a small model and $\alpha \in \mathbb{M}$ satisfies $v(\alpha) > v(m)$ for every $m \in M$, then the sequence $\{nv(\alpha) : n \in \mathbb{N}\}$ is cofinal in the value group of $M\langle\alpha\rangle$, where $M\langle\alpha\rangle$ is a prime model over $M \cup \{\alpha\}$ (i.e. $M\langle\alpha\rangle = \text{dcl}(M \cup \{\alpha\})$).

Lemma 5.3. *Let $p \in S_1(\emptyset)$ be arbitrary.*

- (1) There is at most one global type $\tilde{p} \in S_1(\mathbb{M})$ such that $\tilde{p} \supseteq p \cup \{v(x) > m : m \in \mathbb{M}\}$, and \tilde{p} is \emptyset -invariant.
- (2) Assume \tilde{p} as in (1) exists. Let $M \prec \mathbb{M}$ and $\alpha_1, \dots, \alpha_n \in \mathbb{M}$. Then $(\alpha_1, \dots, \alpha_n)$ realizes $\tilde{p}^{(n)}|M$ if and only if each α_i realizes p , $v(\alpha_1) > v(m)$ for all $m \in M$ and $v(\alpha_{i+1}) > kv(\alpha_i)$ for all $k \in \mathbb{N}$ and $i = 1, \dots, n-1$.

Proof. Part (1) follows from Proposition 5.1 and p -minimality.

Part (2) follows by the same argument as in Lemma 3.4 using Extra Assumptions. \square

Definition 5.4. For an integer $n > 0 \in \mathbb{N}$, we say that a sequence $a_i, i = 1, \dots, L$ of elements \mathbb{Q}_p , is *linearly n -growing* if $v(a_0) > n$ and $v(a_{i+1}) > nv(a_i)$ for all i .

Notice that a subsequence of a linearly n -growing sequence is also linearly n -growing.

Lemma 5.5. For any finite set of formulas $\Delta(x_1, \dots, x_k)$ with parameters from $\tilde{\mathbb{Q}}_p$ there are $n \in \mathbb{N}$ and $d_0 \in \mathbb{N}$ such that any linearly n -growing sequence of elements $a_i \in \mathbb{Q}_p$ of length N contains a Δ -indiscernible subsequence of length at least $\frac{N}{d_0}$.

Proof. Let $\Sigma(x_1, \dots, x_{2k})$ be the partial type that is the union of

$$\Sigma_1 = \left\{ \bigwedge_{1 \leq i < j \leq 2k} P(x_i) \leftrightarrow P(x_j) : P(x) \text{ is an } \mathcal{L}\text{-formula over } \emptyset \right\}$$

and

$$\Sigma_2 = \left\{ (x_1 > n) \wedge \bigwedge_{i=1}^{2k-1} v(x_{i+1}) > nv(x_i) : n \in \mathbb{N} \right\}.$$

Let $(a_i : 1 \leq i \leq N)$ be a sequence of elements in \mathbb{M} having the same type over the empty set. Assume that $v(a_0) > \mathbb{N}$ and $v(a_{i+1}) > nv(a_i)$ for every $i = 1, \dots, N-1$ and $n \in \mathbb{N}$. Then, by Lemma 5.3, the sequence $(a_i : 1 \leq i \leq N)$ realizes $\tilde{p}^{(N)}|_{\mathbb{Q}_p}$ for $p = \text{tp}(a_1/\emptyset)$, and so is indiscernible over \mathbb{Q} by Fact 2.7(2). It follows that

$$\Sigma(x_1, \dots, x_{2k}) \vdash \psi(x_1, \dots, x_k) \leftrightarrow \psi(x_{i_1}, \dots, x_{i_k})$$

for any $1 \leq i_1 < i_2 < \dots < i_k \leq 2k$ and $\psi \in \Delta$.

By compactness, there are finite subsets $\Sigma_1^0 \subseteq \Sigma_1$ and $\Sigma_2^0 \subseteq \Sigma_2$ such that

$$\Sigma_1^0 \cup \Sigma_2^0 \vdash \psi(x_1, \dots, x_k) \leftrightarrow \psi(x_{i_1}, \dots, x_{i_k})$$

for any $1 \leq i_1 < i_2 < \dots < i_k \leq 2k$ and $\psi \in \Delta$.

Let P_1, \dots, P_s be all \mathcal{L} -formulas over \emptyset appearing in Σ_1^0 , and let $n \in \mathbb{N}$ be the largest appearing in Σ_2^0 .

Let $d_0 = 2^s$. Now any linearly n -growing sequence of length N contains a subsequence of length at least $\frac{N}{d_0}$ satisfying the same P_1, \dots, P_s , and this subsequence is Δ -indiscernible. \square

As in the o -minimal case, combining Lemma 3.1 with Lemma 5.5 we can also allow additional parameters in Δ .

Corollary 5.6. For any finite set of \mathcal{L} -formulas $\Delta(x_1, \dots, x_k; y)$ with parameters from $\tilde{\mathbb{Q}}_p$ there are $n \in \mathbb{N}$ and $d \in \mathbb{N}$ such that for all sufficiently large N , for any $c \in \mathbb{Q}_p^{|y|}$, any linearly n -growing sequence $\vec{a} = (a_1, a_2, \dots, a_N)$ of elements from \mathbb{Q}_p contains a $\Delta(x_1, \dots, x_k; c)$ -indiscernible subsequence of length at least $\frac{N}{d}$.

It remains to establish an analog of Fact 3.8 in the p -adic case, demonstrating that there are “enough” linearly n -growing sequences.

For an element $\alpha \in \mathbb{Q}_p$ and $r \in \mathbb{Z}$ we will denote by $B(\alpha, r)$ the closed ball in \mathbb{Q}_p of radius r centered at α , i.e.

$$B(\alpha, r) = \{a \in \mathbb{Q}_p : v(a - \alpha) \geq r\}.$$

Lemma 5.7. *Let $A \subseteq \mathbb{Q}_p$ be a finite non-empty subset with $|A| \geq 2$ and $A \subseteq B(\alpha, r)$ for some $\alpha \in \mathbb{Q}_p$ and $r \in \mathbb{Z}$. Then there is $\alpha' \in B(\alpha, r)$ and $r' > r$ such that $\frac{1}{p}|A| \leq |A \cap B(\alpha', r')| < |A|$.*

Proof. Let $r_1 \in \mathbb{Z}$ be maximal such that some ball $B(\alpha_1, r_1)$ with $\alpha_1 \in B(\alpha, r)$ contains A . The ball $B(\alpha_1, r_1)$ is the union of p balls of radius $r' = r + 1$. Hence there is $\alpha' \in B(\alpha_1, r_1)$ with $\frac{1}{p}|A| \leq |A \cap B(\alpha', r')| < |A|$. \square

Proposition 5.8. *Let $k \in \mathbb{N}$ be positive. For every finite $A \subseteq \mathbb{Q}_p$ with $|A| \geq 2p^{k-1}$ there is $\alpha \in \mathbb{Q}_p$ and elements $a_1, \dots, a_k \in A$ such that the valuations of $\alpha - a_i, i = 1, \dots, k$ are pairwise distinct.*

Proof. Let $A \subseteq \mathbb{Q}_p$ be a finite subset with $|A| \geq 2p^{k-1}$. We set $A_0 = A$, and also choose $\alpha_0 \in \mathbb{Q}_p$ and $r_0 \in \mathbb{Z}$ so that $A \subseteq B(\alpha_0, r_0)$.

Using Lemma 5.7, by induction on $i = 1, \dots, k$ we construct finite sets $A_0 \supsetneq A_1 \supsetneq \dots \supsetneq A_k$, elements $\alpha_1, \dots, \alpha_k \in \mathbb{Q}_p$ and integers $r_1 < r_2 < \dots < r_k$ such that:

- $A_i = B(\alpha_i, r_i) \cap A$;
- $|A_i| \geq \frac{1}{p^i}|A|$;
- $\alpha_i \in B(\alpha_{i-1}, r_{i-1})$ for $i = 1, \dots, k$.

We take $\alpha = \alpha_k$, and for $i = 1, \dots, k$ we let $a_i \in A_{i-1} \setminus A_i$ be arbitrary. Then $a_i \in B(\alpha, r_{i-1}) \setminus B(\alpha, r_i)$, hence $r_{i-1} \leq v(\alpha - a_i) < r_i$. \square

Proposition 5.9. *There are finitely many functions $F_1(x, \bar{y}), \dots, F_s(x, \bar{y})$ definable with parameters from \mathbb{Q}_p such that for any $n \in \mathbb{N}$ there is a constant C such that for any $k \in \mathbb{N}$ the following holds. For any $K \geq 2^{2^{Ck}}$ and any sequence $\vec{a} = (a_1, \dots, a_K)$ in \mathbb{Q}_p there are a linearly n -growing sequence $\vec{b} = (b_1, \dots, b_k)$ of elements in \mathbb{Q}_p , $\vec{c} \in \mathbb{Q}_p^{|\bar{y}|}$ and $i \in \{1, \dots, s\}$ such that one of the sequences*

$$F_i(\vec{b}, \vec{c}) := (F_i(b_1, \vec{c}), F_i(b_2, \vec{c}), \dots, F_i(b_k, \vec{c}))$$

or

$$F_i(\vec{b}, \vec{c}) := (F_i(b_k, \vec{c}), F_i(b_{k-1}, \vec{c}), \dots, F_i(b_1, \vec{c}))$$

is a subsequence of \vec{a} .

Proof. As usual, for a real number R we will denote by $\lceil R \rceil$ the smallest integer N satisfying $N \geq R$.

First, notice that it is sufficient to prove the proposition for $n = 2$. Indeed if a_1, a_2, \dots , is a linearly 2-growing sequence of length N , then for a given n , taking $l = \lceil \log_2 n \rceil$, the sequence a_l, a_{2l}, \dots is a linearly n -growing sequence of length $\frac{N}{l}$.

Assume k is given and $K \geq 2^{2^{Ck}}$, where the constant C can be determined later if needed. Let a_1, \dots, a_K be a sequence of elements of \mathbb{Q}_p .

Case 1. The sequence a_1, \dots, a_K contains at least \sqrt{K} equal elements.

Let's call this element a' . Then we can embed any linearly 2-growing sequence of length $\lceil \sqrt{K} \rceil$ into \vec{a} using the constant map $z \mapsto a'$.

Case 2. The sequence a_1, \dots, a_K does not contain $\lceil \sqrt{K} \rceil$ equal elements. Then it contains at least $K_1 = \sqrt{K} \geq 2^{2^{C_1 k - 1}}$ pairwise distinct elements. Notice that $K_1 \geq 2^{2^{C_1 k}}$ for some constant C_1 depending on C provided $Ck > 1$.

Using Proposition 5.8 we can find an element $\alpha \in \mathbb{Q}_p$ and a subsequence $\vec{a}^1 = (a_1^1, \dots, a_{K_2}^1)$ of \vec{a} with $K_2 \geq \log_p(\frac{1}{2}\sqrt{K_1}) + 1$ such that the valuations $v(\alpha - a_i^1)$ are pairwise distinct for all $1 \leq i \leq K_2$.

Thus, using the map $F(x, \alpha) = x + \alpha$ we can find a sequence $\vec{b} = (b_1, \dots, b_{K_2})$ such that $F(\vec{b}, \alpha)$ is a subsequence of \vec{a} and all of the valuations $v(b_i)$ are pairwise distinct. Notice that the length of \vec{b} is

$$K_2 \geq \frac{1}{\log_2 p} (2^{C_1 k} - 1).$$

It is not hard to compute C_2 from C and p so that $K_2 \geq 2^{C_2 k}$.

By Erdős–Szekeres theorem the sequence \vec{b} contains a subsequence $\vec{b}^1 = (b_1^1, \dots, b_{K_3}^1)$ with $K_3 \geq \lceil \sqrt{K_2} \rceil$ such that the sequence of valuations $(v(b_1^1), \dots, v(b_{K_3}^1))$ is either increasing or decreasing. Using the function $F(x) = x^{-1}$ if needed, we can assume that the sequence is increasing. Again we get $K_3 \geq 2^{C_3 k}$ for an appropriate constant C_3 .

By [6, Lemma 4.1], there is a subsequence $\vec{b}^2 = (b_1^2, \dots, b_{K_4}^2)$ of \vec{b}^1 with $K_4 \geq \frac{1}{2}C_3 k$ such that either for the sequence

$$\vec{v} = (v(b_2^2) - v(b_1^2), v(b_3^2) - v(b_1^2), \dots, v(b_{K_4}^2) - v(b_1^2))$$

we have $v_1 \geq 2, v_{i+1} \geq 2v_i$, or for the sequence

$$\vec{v}' = (v(b_{K_4}^2) - v(b_{K_4-1}^2), v(b_{K_4}^2) - v(b_{K_4-2}^2), \dots, v(b_{K_4}^2) - v(b_1^2))$$

we have $v'_1 \geq 2, v'_{i+1} \geq 2v'_i$.

In the first case, the sequence $\vec{b}^3 = (b_2^2/b_1^2, \dots, b_{K_4}^2/b_1^2)$ is linearly 2-growing and can be embedded into \vec{b}^2 via the transformation $x \mapsto b_1^2 x$.

In the second case, the sequence $\vec{b}^4 = (b_{K_4}^2/b_{K_4-1}^2, \dots, b_{K_4}^2/b_1^2)$ is linearly 2-growing, and its reverse sequence \vec{b}^4 can be embedded into \vec{b}^2 via the transformation $x \mapsto x^{-1}/b_{K_4}^2$. □

Combining Corollary 5.6 and Proposition 5.9 as in the \mathcal{o} -minimal case (see Theorem 3.9), we obtain the main theorem of this section.

Theorem 5.10. *Let $\tilde{\mathbb{Q}}_p$ be a p -minimal expansion of \mathbb{Q}_p satisfying the Extra Assumptions (e.g. the field \mathbb{Q}_p itself). Then for any formula $\varphi(x_1, \dots, x_r; z) \in \mathcal{L}(\mathbb{Q}_p)$, with all x_i singletons, there is a constant $C = C(\varphi)$ such that*

$$R_\varphi^*(n) \leq 2^{2^{Cn}}$$

for all $n \in \mathbb{N}$.

6. RAMSEY GROWTH IN NIP

In this section we consider Ramsey numbers for definable relations of higher arity. We fix a structure \mathcal{M} in a language \mathcal{L} . Following the method of [11] for the semialgebraic case, we obtain the following recursive bound for higher arity Ramsey numbers in arbitrary NIP structures.

Theorem 6.1. *Let \mathcal{M} be an NIP structure, $k \geq 3$ and $\varphi(x_1, \dots, x_k; z) \in \mathcal{L}$ a formula with $|x_1| = \dots = |x_k| = d$. Then for the formula $\psi(x_1, \dots, x_{k-1}; z') = \varphi(x_1, \dots, x_{k-1}; x_k, z)$, taking $m := R_\psi^*(n-1)$, for all large enough n we have*

$$R_\varphi^*(n) \leq 2^{Cm \log m}$$

for some constant $C = C(\varphi)$.

Proof. We are generalizing the argument from [11, Theorem 2.2].

Let $e \in M^{|z|}$ be arbitrary, $\psi(x_1, \dots, x_{k-1}; z') = \varphi(x_1, \dots, x_{k-1}; x_k, z)$, $n \in \mathbb{N}$ large enough, and $m = R_\psi^*(n-1)$. Let $\vec{a} = (a_1, \dots, a_N)$ be a sequence of elements in M^d with $N \geq 2^{Cm \log m}$, where $C = C(\varphi)$ is a constant to be specified later. We need to find a $\varphi(x_1, \dots, x_k; e)$ -indiscernible subsequence of length n .

Let $E \subseteq (M^d)^k$ be the k -ary relation on M^d defined by $\varphi(x_1, \dots, x_k; e)$, i.e. $E = \{(t_1, \dots, t_k) \in (M^d)^k : \mathcal{M} \models \varphi(t_1, \dots, t_k; e)\}$.

The idea is to find a subsequence $\vec{b} = (b_1, \dots, b_{m+1})$ of \vec{a} such that for all $1 \leq i_1 < \dots < i_{k-1} \leq m$, either $(b_{i_1}, \dots, b_{i_{k-1}}, b_i) \in E$ for all $i_{k-1} < i \leq m+1$ or $(b_{i_1}, \dots, b_{i_{k-1}}, b_i) \notin E$ for all $i_{k-1} < i \leq m+1$.

To build a sequence \vec{b} as above, we choose recursively elements b_r in \vec{a} and also subsequences \vec{c}_r of \vec{a} for $r = k-2, k-1, \dots, m+1$ with $\vec{c}_{r+1} \subset \vec{c}_r$ so that the following holds.

- (1) For every $(k-1)$ -subsequence $(b_{i_1}, \dots, b_{i_{k-1}})$ of (b_1, \dots, b_{r-1}) with $i_1 < \dots < i_{k-1}$, either $(b_{i_1}, \dots, b_{i_{k-1}}, b) \in E$ for every $b \in \{b_j : i_{k-1} < j \leq r\} \cup \vec{c}_r$ or $(b_{i_1}, \dots, b_{i_{k-1}}, b) \notin E$ for every $b \in \{b_j : i_{k-1} < j \leq r\} \cup \vec{c}_r$.
- (2) $|\vec{c}_r| \geq \frac{N}{C_1^r C_2^r}$, where C_1, C_2 are some constants depending just on φ .
- (3) Subsequence (b_1, \dots, b_r) appears in \vec{a} before the subsequence \vec{c}_r , i.e. $(b_1, \dots, b_r) \hat{\sim} \vec{c}_r$ is a subsequence of \vec{a} .

We start with $r = k-2$ by taking $(b_1, \dots, b_{k-2}) = (a_1, \dots, a_{k-2})$ and $\vec{c}_{k-2} = (a_{k-1}, \dots, a_N)$. Assume we have obtained (b_1, \dots, b_r) and \vec{c}_r satisfying (1)–(3) above, and we define b_{r+1} and \vec{c}_{r+1} as follows.

Let b_{r+1} be the first element in \vec{c}_r and let \vec{c}_r^* be the sequence \vec{c}_r with the first element removed. Let $\theta(x_k; u)$ be the partitioned formula obtained from $\varphi(x_1, \dots, x_{k-1}, x_k, z)$ by partitioning its variables into two groups x_k and $u = x_1, \dots, x_{k-1}, z$. As the formula θ is NIP, by Fact 2.3 the number of complete $\theta(x_k; u)$ -types over an arbitrary finite set $D \subseteq M^{|z|+(k-1)d}$ of parameters is bounded by $C_3|D|^{C_4}$ for some constants C_3, C_4 depending just on φ .

Let $D = \{(b_{i_1}, \dots, b_{i_{k-1}}, e) : 1 \leq i_1 < \dots < i_{k-1} \leq r+1\}$. Obviously $|D| \leq (r+1)^{k-1}$.

It follows by the pigeonhole principle that there is some complete θ -type $p(x_k) \in S_\theta(D)$ such that the number of elements in \vec{c}_r^* realizing $p(x)$ is at least $\frac{|\vec{c}_r^*|}{C_3|D|^{C_4}} \geq$

$$\frac{|\vec{c}_r^*|-1}{C_3|(r+1)|^{(k-1)C_4}} \geq \frac{|\vec{c}_r^*|}{2C_3|(r+1)|^{(k-1)C_4}}, \text{ provided } |\vec{c}_r^*| \geq 2.$$

We take \vec{c}_{r+1} to be the subsequence of elements of \vec{c}_r realizing p . For $C_1 = 2C_3$ and $C_2 = (k-1)C_4$, using the inductive lower bound for the length of \vec{c}_r and calculating, we obtain $|\vec{c}_{r+1}| \geq \frac{N}{C_1^{(r+1)}(r+1)^{C_2(r+1)}}$, i.e. (2) is satisfied.

Now for any subsequence $(b_{i_1}, \dots, b_{i_{k-1}})$ of (b_1, \dots, b_{r+1}) , we have that either $(b_{i_1}, \dots, b_{i_{k-1}}, b) \in E$ for all $b \in \vec{c}_{r+1}$ or $(b_{i_1}, \dots, b_{i_{k-1}}, b) \notin E$ for all $b \in \vec{c}_{r+1}$. Together with the inductive assumption this implies that (1) is satisfied by (b_1, \dots, b_{r+1}) and \vec{c}_{r+1} . Finally, (3) is clear from the construction.

For \vec{c}_m to be nonempty we need $\frac{N}{C_1^{(r+1)}(r+1)^{C_2(r+1)}} \geq 1$, i.e. $N \geq C_1^m m^{C_2 m}$. It is not hard to find a constant C that depends on C_1, C_2 only so that the condition $N \geq 2^{Cm \log m}$ is sufficient. \square

Remark 6.2. The constant C_4 , in fact, depends just on the VC-density of φ (with a corresponding partition of the variables). By Fact 2.6, in the case of o -minimal theories we can take $C_4 = d$.

By a repeated application of Theorem 6.1 we have an improved bound on Ramsey numbers for relations of higher arities.

Theorem 6.3. *Let \mathcal{M} be an NIP structure, and assume that for all $\psi(x_1, x_2; z)$ we have $R_\psi^*(n) \leq n^c$ for some $c = c(\psi)$ and all n large enough. Then for all $k \geq 3$ and $\varphi(x_1, \dots, x_k; z')$ we have $R_\varphi^*(n) \leq \text{twr}_{k-1}(n^c)$ for some $c = c(\varphi)$ and all n large enough.*

The assumption of Theorem 6.3 is satisfied in distal structures by Fact 1.11.

Corollary 6.4. *Let \mathcal{M} be a distal structure. Then for any $\varphi(x_1, \dots, x_k; z)$ we have $R_\varphi^* \leq \text{twr}_{k-1}(n^c)$ for some $c = c(\varphi)$ and all n large enough.*

It is also satisfied in stable structures by Fact 1.10. We conjecture that it holds in arbitrary NIP structures.

Conjecture 6.5. Let \mathcal{M} be an NIP structure. Then for all formulas $\varphi(x_1, x_2; z)$ we have $R_\varphi^*(n) \leq n^c$ for some $c = c(\varphi)$ and all sufficiently large n .

This conjecture is equivalent to saying that the Erdős-Hajnal conjecture holds for all graphs definable in NIP structures. We refer to [9, 10] for a discussion.

REFERENCES

- [1] Hans Adler, *An introduction to theories without the independence property*, Archive for Mathematical Logic **5** (2008).
- [2] Noga Alon, János Pach, Rom Pinchasi, Rados Radoičić, and Micha Sharir, *Crossing patterns of semi-algebraic sets*, Journal of Combinatorial Theory, Series A **111** (2005), no. 2, 310–326.
- [3] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko, *Vapnik-Chervonenkis density in some theories without the independence property, I*, Trans. Amer. Math. Soc. **368** (2016), no. 8, 5889–5949. MR3458402
- [4] Saugata Basu, *Combinatorial complexity in o -minimal geometry*, Proceedings of the London Mathematical Society **100** (2010), no. 2, 405–428.
- [5] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U. S. A. **32** (1946), 331–332. MR0018694
- [6] Boris Bukh and Jirí Matousek, *Erdős-szekeres-type statements: Ramsey function and decidability in dimension 1*, Duke Mathematical Journal **163** (2014), no. 12, 2243–2270.
- [7] Artem Chernikov and Pierre Simon, *Externally definable sets and dependent pairs II*, Transactions of the American Mathematical Society **367** (2015), no. 7, 5217–5235.
- [8] Artem Chernikov and Sergei Starchenko, *A note on the Erdős-Hajnal property for stable graphs*, Preprint, arXiv:1504.08252 (2015).
- [9] ———, *Regularity lemma for distal structures*, Journal of the European Mathematical Society, to appear (arXiv:1507.01482) (2015).
- [10] ———, *Definable regularity lemmas for NIP hypergraphs*, Preprint, arXiv:1607.07701 (2016).
- [11] David Conlon, Jacob Fox, János Pach, Benny Sudakov, and Andrew Suk, *Ramsey-type results for semi-algebraic relations*, Transactions of the American Mathematical Society **366** (2014), no. 9, 5043–5065.
- [12] David Conlon, Jacob Fox, and Benny Sudakov, *Hypergraph Ramsey numbers*, Journal of the American Mathematical Society **23** (2010), no. 1, 247–266.

- [13] Marek Elias, Jirí Matousek, Edgardo Roldán-Pensado, and Zuzana Safernová, *Lower bounds on geometric Ramsey functions*, SIAM Journal on Discrete Mathematics **28** (2014), no. 4, 1960–1970.
- [14] Doug Ensley and Rami Grossberg, *Ramsey’s theorem in stable structures*, manuscript, 1997.
- [15] Paul Erdős, *Some remarks on the theory of graphs*, Bulletin of the American Mathematical Society **53** (1947), no. 4, 292–294.
- [16] Paul Erdős, András Hajnal, and Richard Rado, *Partition relations for cardinal numbers*, Acta Mathematica Hungarica **16** (1965), no. 1-2, 93–196.
- [17] Paul Erdős and Richard Rado, *Combinatorial theorems on classifications of subsets of a given set*, Proceedings of the London mathematical Society **3** (1952), no. 1, 417–439.
- [18] Paul Erdős and George Szekeres, *A combinatorial problem in geometry*, Compositio Mathematica **2** (1935), 463–470.
- [19] Ronald Graham, Bruce Rothschild, and Joel Spencer, *Ramsey theory*, Vol. 20, John Wiley & Sons, 1990.
- [20] Vincent Guingona, *On uniform definability of types over finite sets*, The Journal of Symbolic Logic **77** (2012), no. 02, 499–514.
- [21] Deirdre Haskell and Dugald Macpherson, *A version of o-minimality for the p-adics*, The Journal of Symbolic Logic **62** (1997), no. 04, 1075–1092.
- [22] Ehud Hrushovski and Anand Pillay, *On NIP and invariant measures*, Journal of the European Mathematical Society **13** (2011), no. 4, 1005–1061.
- [23] Hunter Johnson and Michael Laskowski, *Compression schemes, stable definable families, and o-minimal structures*, Discrete & Computational Geometry **43** (2010), no. 4, 914–926.
- [24] Itay Kaplan and Saharon Shelah, *A dependent theory with few indiscernibles*, Israel Journal of Mathematics **202** (2014), no. 1, 59–103.
- [25] Michael Laskowski, *Vapnik-Chervonenkis classes of definable sets*, Journal of the London Mathematical Society **2** (1992), no. 2, 377–384.
- [26] Angus Macintyre, *On definable subsets of p-adic fields*, The Journal of Symbolic Logic **41** (1976), no. 3, 605–610.
- [27] Maryanthe Malliaris and Saharon Shelah, *Regularity lemmas for stable graphs*, Transactions of the American Mathematical Society **366** (2014), no. 3, 1551–1585.
- [28] Jirí Matousek, *Lectures on discrete geometry*, Vol. 108.
- [29] Chris Miller, *Exponentiation is hard to avoid*, Proceedings of the American Mathematical Society **122** (1994), no. 1, 257–259.
- [30] Chris Miller and Sergei Starchenko, *A growth dichotomy for o-minimal expansions of ordered groups*, Trans. Amer. Math. Soc. **350** (1998), no. 9, 3505–3521. MR1491870
- [31] Shay Moran and Amir Yehudayoff, *Sample compression schemes for VC classes*, Preprint, arXiv:1503.06960 (2015).
- [32] FP Ramsey, *On a problem of formal logic*, Proceedings of the London Mathematical Society **2** (1930), no. 1, 264–286.
- [33] Thomas Scanlon, *O-minimality as an approach to the André-Oort conjecture*, Panoramas et Synthèses, to appear.
- [34] S. Shelah, *Classification theory and the number of nonisomorphic models*, Second, Studies in Logic and the Foundations of Mathematics, vol. 92, North-Holland Publishing Co., Amsterdam, 1990.
- [35] Saharon Shelah, *Around classification theory of models*, Springer, 1986.
- [36] ———, *Strongly dependent theories*, Israel Journal of Mathematics **204** (2014), no. 1, 1–83.
- [37] Pierre Simon, *A guide to NIP theories*, Cambridge University Press, 2015.
- [38] Lou Van den Dries, *Tame topology and o-minimal structures*, Vol. 248, Cambridge university press, 1998.
- [39] Alex J Wilkie, *Model completeness results for expansions of the ordered field of real numbers by restricted pfaffian functions and the exponential function*, Journal of the American Mathematical Society **9** (1996), no. 4, 1051–1094.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA LOS ANGELES, LOS ANGELES, CA
90095-1555

E-mail address: `chernikov@math.ucla.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NOTRE DAME, NOTRE DAME, IN 46556

E-mail address: `Starchenko.1@nd.edu`

ZUKUNFTSKOLLEG, DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF KONSTANZ,
BOX 216, 78457 KONSTANZ, GERMANY

E-mail address: `margaret.thomas@wolfson.oxon.org`